SCORE Search Results Details for Application 10578781 and Search Result 20081104_154454_us-10-578-781-1.rng.

Score Home Page | Retrieve Application List | SCORE System Overview | SCORE FAQ | Comments / Suggestions

This page gives you Search Results detail for the Application 10578781 and Search Result 20081104_154454_us-10-578-781-1.rng.

Go Back to previous page

GenCore version 6.3

OM nucleic - nucleic search, using sw model

Run on: November 4, 2008, 17:10:32 ; Search time 243 Seconds
(without alignments)
44258.760 Million cell updates/sec

Title: US-10-578-781-1
Perfect score: 756
Sequence: 1 atgggtgccgatatcaaaaa.....aggagcaaattgaacattag 756

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 11806651 seqs, 7113014948 residues

Total number of hits satisfying chosen parameters: 23613302

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_200808:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000:*
4: geneseqn2001a:*
5: geneseqn2001b:*
6: geneseqn2002a:*
7: geneseqn2002b:*
8: geneseqn2003a:*
9: geneseqn2003b:*
10: geneseqn2003c:*
11: geneseqn2003d:*
12: geneseqn2004a:*
13: geneseqn2004b:*
14: geneseqn2004c:*
15: geneseqn2004d:*
16: geneseqn2004e:*
17: geneseqn2004f:*
18: geneseqn2005a:*
19: geneseqn2005b:*
20: geneseqn2005c:*
21: geneseqn2006a:*

22: geneseqn2006b:*
23: geneseqn2006c:*
24: geneseqn2006d:*
25: geneseqn2007a:*
26: geneseqn2007b:*
27: geneseqn2007c:*
28: geneseqn2007d:*
29: geneseqn2008:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	756	100.0	756	18	AEA00728	Aea00728 Brevibaci
2	273.4	36.2	665	6	ABK78469	Abk78469 Bacillus
3	192.2	25.4	1438	12	ADN60500	Adn60500 B. lichen
4	178.2	23.6	777	21	AEH93992	Aeh93992 SigG codi
5	173.8	23.0	381	6	ABK74144	Abk74144 Bacillus
6	147	19.4	631	6	ABK74048	Abk74048 Bacillus
7	78	10.3	1110	12	ADH97013	Adh97013 S. pneumo
8	78	10.3	1110	18	AEC13368	Aec13368 Streptoco
9	76.4	10.1	915	12	ADK44581	Adk44581 Streptoco
10	76.4	10.1	915	21	AEJ68509	Aej68509 Streptoco
11	76.4	10.1	915	21	AEJ75484	Aej75484 Streptoco
12	76.4	10.1	915	21	AEJ82844	Aej82844 Streptoco
13	76.4	10.1	915	21	AEL05163	Ael05163 Streptoco
14	76.4	10.1	915	21	AEL12413	Ael12413 Streptoco
15	76.4	10.1	915	21	AEL50821	Ael50821 Streptoco
16	76.4	10.1	915	25	AEM07844	Aem07844 Streptoco
17	76.4	10.1	915	25	AEM66066	Aem66066 Streptoco
18	76.4	10.1	915	25	AEM86645	Aem86645 Streptoco
19	76.4	10.1	915	25	AGI20663	Agi20663 Streptoco
20	76.4	10.1	915	25	AEN48202	Aen48202 Streptoco
21	76.4	10.1	915	25	AEN55537	Aen55537 Streptoco
22	76.4	10.1	915	25	AEN40434	Aen40434 Streptoco
23	76.4	10.1	915	25	AGI76906	Agi76906 Streptoco
24	76.4	10.1	915	25	AEN08741	Aen08741 Streptoco
25	76.4	10.1	915	25	AGV09876	Agv09876 Streptoco
26	76.4	10.1	915	25	AGV21123	Agv21123 Streptoco
27	76.4	10.1	915	25	AJE78366	Aje78366 Streptoco
28	76.4	10.1	915	25	AJE70154	Aje70154 Streptoco
29	76.4	10.1	915	25	AJE86340	Aje86340 Streptoco
30	76.4	10.1	915	25	AJE95472	Aje95472 Streptoco
31	76.4	10.1	915	25	AJE61089	Aje61089 Streptoco
32	76.4	10.1	915	25	AGV40930	Agv40930 Streptoco
33	76.4	10.1	915	25	AGV46355	Agv46355 Streptoco
34	76.4	10.1	915	25	AJF01903	Ajf01903 Streptoco
35	76.4	10.1	915	25	AJF07993	Ajf07993 Streptoco
36	76.4	10.1	915	25	AJF13317	Ajf13317 Streptoco
37	76.4	10.1	915	25	AJF53284	Ajf53284 Streptoco
38	76.4	10.1	915	25	AJF18960	Ajf18960 Streptoco
39	76.4	10.1	915	25	AJG97788	Ajg97788 Streptoco
40	76.4	10.1	915	25	ALK14103	Alk14103 S.pneumon
41	76.4	10.1	915	25	ALT08207	Alt08207 Streptoco
42	76.4	10.1	915	25	ANK69294	Ank69294 Streptoco
43	76.4	10.1	915	25	ANJ76477	Anj76477 S. pneumo
44	76.4	10.1	915	25	ANK74634	Ank74634 Streptoco
45	76.4	10.1	915	26	ANN03548	Ann03548 S. pneumo

ALIGNMENTS

RESULT 1

```
AEA00728
ID   AEA00728 standard; DNA; 756 BP.
XX
AC   AEA00728;
XX
DT   28-JUL-2005  (first entry)
XX
DE   Brevibacillus choshinensis DNA #1.
XX
KW   Cell culture; Brevibacillus choshinensis; gene; ds.
XX
OS   Brevibacillus choshinensis.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..756
FT                   /*tag= a
FT                   /product= "B. choshinensis protein #1"
XX
PN   WO2005045005-A1.
XX
PD   19-MAY-2005.
XX
PF   08-NOV-2004; 2004WO-JP016912.
XX
PR   11-NOV-2003; 2003JP-00381606.
XX
PA   (HGET ) HIGETA SHOYU KK.
XX
PI   Hanagata H, Nishijyo T;
XX
DR   WPI; 2005-366840/37.
DR   P-PSDB; AEA00729.
XX
PT   New Brevibacillus choshinensis, that does not form spores and which shows
PT   low extracellular or intracellular protease activity, useful as host for
PT   producing recombinant protein.
XX
PS   Claim 4; SEQ ID NO 1; 103pp; Japanese.
XX
CC   The invention relates to a Brevibacillus choshinensis HPD31-SP3 (FERMBP-
CC   08479), which does not form spores and which has mycological
CC   characteristics such as cell size and rod shape and physiological
CC   characteristics such as negative for nitrate reduction and positive for
CC   citric acid utilization, oxidase and catalase, and showing low
CC   extracellular protease activity. The invention also relates to a
CC   transformed B. choshinensis using a vector containing the gene encoding
CC   the protein of the invention. B. choshinensis is useful as a host for
CC   producing a recombinant protein and for producing a protein by culturing
CC   a transformed host. It has decreased extracellular protein degradation
CC   activity when compared with other strains. This sequence represents B.
CC   choshinensis DNA of the invention.
XX
SQ   Sequence 756 BP; 207 A; 166 C; 211 G; 172 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 756;  DB 18;  Length 756;
  Best Local Similarity   100.0%;  Pred. No. 3e-234;
  Matches  756;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 ATGGGTGCCGATATCAAAAATGCGAGTCAACCATTTCTGACCAATGACCAAGTGAAAGAT 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 ATGGGTGCCGATATCAAAAATGCGAGTCAACCATTTCTGACCAATGACCAAGTGAAAGAT 60

Qy         61 TTGATAGCCAAGAGCCAAGCTGGCGATACGGATGCACGTGAGCTTCTCGTGAATAGCAAT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 TTGATAGCCAAGAGCCAAGCTGGCGATACGGATGCACGTGAGCTTCTCGTGAATAGCAAT 120

Qy        121 ATCAGACTGGTCTGGTCCGTCGTCCAGCGCTTTATCAACCGCGGGTATGAAGCGGATGAT 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ATCAGACTGGTCTGGTCCGTCGTCCAGCGCTTTATCAACCGCGGGTATGAAGCGGATGAT 180
```

```
Qy          181 TTGTTTCAGATCGGTTGCATTGGCTTGCTCAAGGCCGTTGACAAGTTCGATCTTTCGTAC 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          181 TTGTTTCAGATCGGTTGCATTGGCTTGCTCAAGGCCGTTGACAAGTTCGATCTTTCGTAC 240

Qy          241 GATGTGAGATTTTCGACCTATGCGGTGCCAATGATCATCGGAGAAATTCAACGCTTTTTG 300
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          241 GATGTGAGATTTTCGACCTATGCGGTGCCAATGATCATCGGAGAAATTCAACGCTTTTTG 300

Qy          301 CGCGATGACGGTACGGTTAAGGTCAGTCGATCGTTAAAAGAAACAGCGAATAAGGTGCGG 360
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          301 CGCGATGACGGTACGGTTAAGGTCAGTCGATCGTTAAAAGAAACAGCGAATAAGGTGCGG 360

Qy          361 CGATCAAAGGATGAATTGTACAAGCAATTCGGCCGTGCCCCCACGATCGCAGAAGTGGCA 420
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          361 CGATCAAAGGATGAATTGTACAAGCAATTCGGCCGTGCCCCCACGATCGCAGAAGTGGCA 420

Qy          421 GAAGCAGTGGGAATCACGCCGGAGGAAGTAGTCTTTGCGCAAGAGGCAAGCAGAGCGCCT 480
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          421 GAAGCAGTGGGAATCACGCCGGAGGAAGTAGTCTTTGCGCAAGAGGCAAGCAGAGCGCCT 480

Qy          481 TCCTCCATCCATGAGACCGTTTTTTGAAAATGACGGCGATCCCATCACACTGATCGATCAG 540
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          481 TCCTCCATCCATGAGACCGTTTTTTGAAAATGACGGCGATCCCATCACACTGATCGATCAG 540

Qy          541 ATAGCGGATGAAGGTGTGAACAAGTGGTTTGAGAAAATTGCCTTGAAGGACGCCATCAGC 600
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          541 ATAGCGGATGAAGGTGTGAACAAGTGGTTTGAGAAAATTGCCTTGAAGGACGCCATCAGC 600

Qy          601 AGGCTGAGCGAGCGTGAGCAGCTCATCGTCTACCTGCGCTATTACAAGGATCAGACACAG 660
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          601 AGGCTGAGCGAGCGTGAGCAGCTCATCGTCTACCTGCGCTATTACAAGGATCAGACACAG 660

Qy          661 TCTGAGGTAGCAGAGCGTCTAGGGATTTCGCAGGTCCAGGTCTCGCGTCTGGAAAAGCGT 720
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          661 TCTGAGGTAGCAGAGCGTCTAGGGATTTCGCAGGTCCAGGTCTCGCGTCTGGAAAAGCGT 720

Qy          721 ATCCTGCTAACGATCAAGGAGCAAATTGAACATTAG 756
                |||||||||||||||||||||||||||||||||||
Db          721 ATCCTGCTAACGATCAAGGAGCAAATTGAACATTAG 756
```

```
RESULT 2
ABK78469
ID   ABK78469 standard; DNA; 665 BP.
XX
AC   ABK78469;
XX
DT   13-AUG-2002 (first entry)
XX
DE   Bacillus clausii genomic sequence tag (GST) #1312.
XX
KW   Differential gene expression; genomic sequenced tag; GST;
KW   altered culture condition; environmental stress;
KW   physiological provocation; ds.
XX
OS   Bacillus clausii.
XX
PN   WO200229113-A2.
XX
PD   11-APR-2002.
XX
PF   05-OCT-2001; 2001WO-US031437.
XX
PR   06-OCT-2000; 2000US-00680598.
PR   27-MAR-2001; 2001US-0279526P.
XX
PA   (NOVO) NOVOZYMES BIOTECH INC.
PA   (NOVO) NOVOZYMES AS.
```

XX
PI Berka R, Clausen IG;
XX
DR WPI; 2002-416684/44.
XX
PT Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second Bacillus
PT cells, by using substrate containing Bacillus genomic sequenced tag
PT array.
XX
PS Claim 11; SEQ ID NO 5760; 200pp; English.
XX
CC The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive follow
CC -up characterisation is unnecessary, when one spot on an array equals one
CC gene or one open reading frame, since sequence information is available.
CC This sequence represents a genomic sequence tag (GST) used in the method
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 665 BP; 210 A; 134 C; 159 G; 161 T; 0 U; 1 Other;

Query Match 36.2%; Score 273.4; DB 6; Length 665;
Best Local Similarity 67.4%; Pred. No. 1.9e-77;
Matches 399; Conservative 0; Mismatches 192; Indels 1; Gaps 1;

```
Qy          37 CTGACCAATGACCAAGTGAAAGATTTGATAGCCAAGAGCCAAGCTGGCGATACGGATGCA 96
               ||  || || | ||||||||||||  | || || || ||||| |  ||||| || || |||
Db          40 CTATCCGATAAACAAGTGAAAGAGCTTATTGCAAAAAGCCAGGAAGGCGACACAGAAGCA 99

Qy          97 CGTGAGCTTCTCGTGAATAGCAATATCAGACTGGTCTGGTCCGTCGTCCAGCGCTTTATC 156
               ||  ||      |||| ||     || |   | || |||||||| || || || || ||| |
Db         100 CGGGATTCGATCGTCAACCATAACACACGTCTCGTCTGGTCAGTGGTTCAACGTTTTTTG 159

Qy         157 AACCGCGGGTATGAAGCGGATGATTTGTTTCAGATCGGTTGCATTGGCTTGCTCAAGGCC 216
               ||||||  |||||  || |||||  |  |||||  || || |||||||||  ||  |  |||| |
Db         160 AATCGCGGTTATGAGGCAGATGACCTTTTTCAAATTGGCTGCATTGGTTTAATTAAGTCT 219

Qy         217 GTTGACAAGTTCGATCTTTCGTACGATGTGAGATTTTCGACCTATGCGGTGCCAATGATC 276
               ||  |||||  ||  || |||||  |||||  ||||  |||||| || ||||||  ||||||  |||||
Db         220 GTCGACAAATTTGACCTTTCCTACGACGTGAAATTTTCCACGTATGCTGTGCCGATGATT 279

Qy         277 ATCGGAGAAATTCAACGCTTTTTGCGCGATGACGGTACGGTTAAGGTCAGTCGATCGTTA 336
               ||  ||  |||||  |||||  |||  ||||  |||||  ||  ||  ||  ||  ||  ||  ||   |
Db         280 ATTGGTGAAATCCAACGGTTTCTGCGGGATGATGGCACAGTGAAAGTAAGCCGGTCCATT 339

Qy         337 AAAGAAACAGCGAATAAGGTGCGGCGATCAAAGGATGAATTGTACAAGCAATTCGGCCGT 396
               ||||||  |   |||||  |  ||   |  ||||  ||  |||  ||   ||      |  ||||
Db         340 AAAGAATTAAGCAATAAAATCCGCAAAGCAAAAGACGAACTGACGAAAACGCTGCGCCGG 399

Qy         397 GCCCCCACGATCGCAGAAGTGGCAGAAGCAGTGGGAATCACGCCGGAGGAAGTAGTCTTT 456
               ||  ||  ||  ||     ||   |  ||  |||     |  ||   |  ||||| ||||||  |  |||  |||
Db         400 GCACCGACCATTAATGAGATCGCTGAACATTTAGGCGTGACGCCTGAGGAAATTGTATTT 459

Qy         457 GCGCAAGAGGCAAGCAGAGCGCCTTCCTCCATCCATGAGACCGTTTTTTGAAAATGACGGC 516
```

```
                ||    ||| || | | |       |||||| |||||||| || |||| ||||||||||||||
Db          460 GCTGGAGATGCCAACCGGAGCTTGTCCTCAATCCATGAAACGGTTTATGAAAATGACGGC 519

Qy          517 GATCCCATCACACTGATCGATCAGATAGCGGATGAAGGTGTGAACAAGTGGTTTGAGAAA 576
                ||||| || ||||| | ||||| || || ||  |          ||| |||||||| ||
Db          520 GATCCGATTACACTTCTAGATCAAATTGCCGACCACTCACAAGTCAAATGGTTTGACAAG 579

Qy          577 ATTGCCTTGAAGGACGCCATCAGCAGGCTGAGCGAGCGTGAGCAGCTCATCG 628
                |||| ||| || || || ||  ||   ||  ||||  | |||| ||| || |
Db          580 ATTG-CTTTAAAGAAGCGATTTGCGACCTTGGCGANAGGGAGCGGCTAATTG 630


RESULT 3
ADN60500
ID   ADN60500 standard; DNA; 1438 BP.
XX
AC   ADN60500;
XX
DT   01-JUL-2004 (first entry)
XX
DE   B. licheniformis sporulation related polynucleotide, seq id 172.
XX
KW   Mutant host cell; sporulation; oxidoreductase; transferase; hydrolase;
KW   lyase; isomerase; ligase; gene; ds.
XX
OS   Bacillus licheniformis.
XX
PN   WO2003087148-A2.
XX
PD   23-OCT-2003.
XX
PF   25-MAR-2003; 2003WO-DK000200.
XX
PR   10-APR-2002; 2002DK-00000533.
XX
PA   (NOVO) NOVOZYMES AS.
XX
PI   Andersen JT, Jorgensen ST, Rasmussen MD, Olsen PB, Clausen IG;
XX
DR   WPI; 2004-122131/12.
DR   P-PSDB; ADN60501.
XX
PT   A Bacillus licheniformis mutant host cell for producing a product of
PT   interest e.g. vitamins, antibiotics and enzymes.
XX
PS   Claim 1; SEQ ID NO 172; 319pp; English.
XX
CC   The invention relates to a Bacillus licheniformis mutant host cell
CC   derived from a parent B. licheniformis host cell. The mutant host cell is
CC   mutated in one or more genes encoding one or more polypeptides involved
CC   in sporulation. The host cell comprises one or more heterologous genes
CC   present in at least two copies, encoding one or more heterologous
CC   polypeptides. The heterologous genes are stably integrated into the
CC   genome of the cell without leaving any antibiotic resistance marker genes
CC   at the site of integration. The heterologous genes are transcribed from a
CC   heterologous promoter or from an artificial promoter, and are comprised
CC   in an operon, preferably a polycistronic operon. The heterologous
CC   polypeptide is an antimicrobial peptide, or a fusion peptide comprising a
CC   peptide part which in its native form has antimicrobial activity. The
CC   heterologous polypeptide is an enzyme, preferably a secreted enzyme. The
CC   enzyme is an enzyme of a class selected from the group of enzyme classes
CC   consisting of oxidoreductases (EC 1), transferases (EC 2), hydrolases (EC
CC   3), lyases (EC 4), isomerases (EC 5), and ligases (EC 6). The Bacillus
CC   licheniformis is useful in a process for producing at least one product
CC   of interest, comprising cultivating a B. licheniformis mutant host cell
CC   in a suitable medium, whereby the said product is produced. The process
CC   further comprises isolating or purifying the product of interest. The
CC   current sequence represents a B. licheniformis sporulation related
CC   polynucleotide.
```

```
XX
SQ   Sequence 1438 BP;  421 A;  324 C;  355 G;  338 T;  0 U;  0 Other;

  Query Match             25.4%;  Score 192.2;  DB 12;  Length 1438;
  Best Local Similarity   63.7%;  Pred. No. 7.1e-51;
  Matches  309;  Conservative    0;  Mismatches  173;  Indels    3;  Gaps    1;

Qy          1 ATGGGTGCCGATATCAAAAATGCGAGTCAACCATTTCTGACCAATGACC---AAGTGAAA 57
              |||| ||  ||  | ||||| |  |  ||        || |   || ||||   ||||||||
Db        954 ATGGATGTGGAGGTTAAAAAAGAAAACCAGAACACTCAGCTTAAAGACCATGAAGTGAAA 1013

Qy         58 GATTTGATAGCCAAGAGCCAAGCTGGCGATACGGATGCACGTGAGCTTCTCGTGAATAGC 117
              ||  ||||    || |||||     ||||||    | ||| | || || ||| |  | |
Db       1014 GAACTGATTAAAAACAGCCAGAACGGCGATCAAAAAGCAAGGGACCTCCTCATAGAAAAA 1073

Qy        118 AATATCAGACTGGTCTGGTCCGTCGTCCAGCGCTTTATCAACCGCGGGTATGAAGCGGAT 177
              || ||  | || || ||||| ||||| ||||| ||| | ||| | || |||||  | ||
Db       1074 AACATGCGTCTTGTTTGGTCTGTCGTTCAGCGTTTTTTGAACAGAGGCTATGAGCCTGAC 1133

Qy        178 GATTTGTTTCAGATCGGTTGCATTGGCTTGCTCAAGGCCGTTGACAAGTTCGATCTTTCG 237
              ||  | ||||| ||||| ||||| ||| |  | ||| | || ||||| ||||||||||
Db       1134 GACCTCTTTCAAATCGGCTGCATCGGCCTCTTGAAGTCGGTGGACAAATTCGATCTTTCC 1193

Qy        238 TACGATGTGAGATTTTCGACCTATGCGGTGCCAATGATCATCGGAGAAATTCAACGCTTT 297
              || || ||  | ||||| ||||| || || || |||||| ||||||| ||||| || |||
Db       1194 TATGACGTTCGGTTTTCCACCTACGCCGTTCCGATGATTATCGGCGAGATTCAGCGGTTT 1253

Qy        298 TTGCGCGATGACGGTACGGTTAAGGTCAGTCGATCGTTAAAAGAAACAGCGAATAAGGTG 357
               |  | |||||||| || || || || || || ||| | |||||| |  || ||  |
Db       1254 ATCAGAGATGACGGAACCGTCAAAGTGAGCCGCTCGCTGAAAGAACTCGGCAACAAAATC 1313

Qy        358 CGGCGATCAAAGGATGAATTGTACAAGCAATTCGGCCGTGCCCCCACGATCGCAGAAGTG 417
              ||||| | || || ||  | | |||| |  |||||| |  || || |  ||| |
Db       1314 CGGCGGGCGAAAGACGAGCTTTCCAAGTCAAACGGCCGGATTCCGACCGTTCAGGAAATC 1373

Qy        418 GCAGAAGCAGTGGGAATCACGCCGGAGGAAGTAGTCTTTGCGCAAGAGGCAAGCAGAGCG 477
              || ||    | | |||||  | || || || ||  | || || || ||  | | |
Db       1374 GCCGATTATCTCGAAATCAGTTCAGAAGAGGTCGTGATGGCCCAGGAAGCGGTCCGCTCT 1433

Qy        478 CCTTC 482
              || ||
Db       1434 CCCTC 1438


RESULT 4
AEH93992
ID   AEH93992 standard; cDNA; 777 BP.
XX
AC   AEH93992;
XX
DT   27-JUL-2006  (first entry)
XX
DE   SigG coding sequence.
XX
KW   ss; gene; protein production; sigma factor; RNA polymerase;
KW   alkaline protease; food.
XX
OS   Bacillus sp.; KSM-9865.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..777
FT                   /*tag= a
FT                   /product= "SigG"
XX
PN   JP2006136221-A.
XX
PD   01-JUN-2006.
XX
PF   10-NOV-2004; 2004JP-00326973.
```

XX
PR 10-NOV-2004; 2004JP-00326973.
XX
PA (KAOS) KAO CORP.
XX
PI Sumitomo N, Okuda T, Takimura Y, Sato T, Kobayashi T;
XX
DR WPI; 2006-385028/40.
DR P-PSDB; AEH93994.
XX
PT Novel sporulation related gene encoding SigE protein or SigG protein
PT having alkaline protease activity, useful in foodstuffs such as alcoholic
PT beverage, bean paste, soy sauce, pharmaceuticals and cosmetics.
XX
PS Claim 1; SEQ ID NO 2; 15pp; Japanese.
XX
CC This sequence represents the SigG sequence which encodes the sigma factor
CC which is a subunit of RNA polymerase. A microorganism transformed with
CC the SigE or SigG sequence is useful for producing a protein or
CC polypeptide having alkaline protease activity. The sigma peptides are
CC useful in foodstuffs e.g. alcoholic beverage, bean paste, soy sauce,
CC pharmaceuticals and cosmetics. The protein having alkaline protease
CC activity can be produced efficiently using non-sporulated microorganisms.
XX
SQ Sequence 777 BP; 264 A; 123 C; 189 G; 201 T; 0 U; 0 Other;

Query Match 23.6%; Score 178.2; DB 21; Length 777;
Best Local Similarity 53.9%; Pred. No. 1.8e-46;
Matches 391; Conservative 0; Mismatches 328; Indels 6; Gaps 1;

```
Qy          31 CCATTTCTGACCAATGACCAAGTGAAAGATTTGATAGCCAAGAGCCAAGCTGGCGATACG 90
               ||  ||  |||  |||||  ||  ||     |  ||  |     |  |  ||||  ||  ||
Db          49 CCTGTTTTGAAGAATGAAGAAATGCGGAAGTTATTTCGTGAAATGCAAGGCGGAGAACTT 108

Qy          91 GATGCACGTGAGCTTCTCGTGAATAGCAATATCAGACTGGTCTGGTCCGTCGTCCAGCGC 150
                ||||  |  ||     |||||||||  |  ||     ||  ||  ||     ||  |  ||  ||
Db         109 TCTGCAAGAGAAAAGCTCGTGAATGGGAACCAAAGGCTCGTTCTAAGTGTTATTCAACGG 168

Qy         151 TTTATCAACCGCGGGTATGAAGCGGATGATTTGTTTCAGATCGGTTGCATTGGCTTGCTC 210
               ||||  |||||||  ||  |     |  ||||||  ||  ||||||  |  ||  ||||  ||  ||  |
Db         169 TTTAACAACCGTGGTGAATTTGTAGATGACTTATTTCAAGTAGGCTGCATCGGGTTAATG 228

Qy         211 AAGGCCGTTGACAAGTTCGATCTTTCGTACGATGTGAGATTTTCGACCTATGCGGTGCCA 270
               |||  |  ||||  ||  ||  ||  ||     |  |||||  |  ||||||  ||  ||||||  ||  |||
Db         229 AAGTCGATTGATAATTTTGACCTGGGTCAGAATGTTAAGTTTTCTACATATGCAGTACCA 288

Qy         271 ATGATCATCGGAGAAATTCAACGCTTTTTGCGCGATGACGGTACGGTTAAGGTCAGTCGA 330
               |||||  ||||||  ||  ||     ||||  |  |  ||  |||  |     |  ||  |  |||     ||
Db         289 ATGATAATCGGGGAGATAAGGCGCTATCTACGAGATAATAATCCGATCAGGGTATCCCGC 348

Qy         331 TCGTTAAAAGAAACAGCGAATAAGGTGCGGCGATCAAAGGATGAATTGTACAAGCAATTC 390
               ||  ||     ||  |  ||  |  ||||  |  ||     |  |||     ||  |  ||
Db         349 TCATTGCGTGATATTGCCTACAAGGCCCTGCAGGTCAGGGAAAGGCTGATGAGTGAAACA 408

Qy         391 GGCCGTGCCCCCACGATCGCAGAAGTGGCAGAAGCAGTGGGAATCACGCCGGAGGAAGTA 450
                  |  |  ||  ||     |  |||||  |  ||  ||||  |  |  |  |  |||  |||||  |
Db         409 TCAAGGGAGCCTACCGCAGAAGAAATTTCAAAAGTACTTGAAGTACCGCATGAGGAGATT 468

Qy         451 GTCTTTGCGCAAGAGGCAAGCAGAGCGCCTTCCTCCATCCATGAGACCGTTTTTGAAAAT 510
               ||  |||||     |  |  |||  |     |  ||     |||  |     ||||  |  |  |  |  ||
Db         469 GTTTTTGCTTTAAATGCTATTCAGGATCCCGTGTCCTTATTTGAGCCTATCTATAATGAT 528

Qy         511 GACGGCGATCCCATCACACTGATCGATCAGATAGCGGATGAAGGTGTGAA------CAAG 564
               |  ||  |||||  ||     |  |  |||||  ||     ||  |||     ||     |  ||
Db         529 GGTGGAGATCCGATTTATGTATTGGATCAAATCAGTGACGAAAAGAACAAAGATATCCAA 588

Qy         565 TGGTTTGAGAAAATTGCCTTGAAGGACGCCATCAGCAGGCTGAGCGAGCGTGAGCAGCTC 624
               |||  |  ||  |  ||  ||  |  ||  ||  |  ||  |  ||  |  ||  |  |||||
Db         589 TGGATAGAAGAGATAGCACTAAAAGAAGGTATGAGACGTCTCAATGACAGGGAAAAGCTC 648
```

```
Qy          625 ATCGTCTACCTGCGCTATTACAAGGATCAGACACAGTCTGAGGTAGCAGAGCGTCTAGGG 684
                ||  |      ||  |  ||   |  |   |  ||  ||     ||  ||||||  ||      |  |||
Db          649 ATTTTAAGAAAACGGTTTTTTCAAGGGAAAACGCAAATGGAAGTAGCTGATGAAATCGGG 708

Qy          685 ATTTCGCAGGTCCAGGTCTCGCGTCTGGAAAAGCGTATCCTGCTAACGATCAAGGAGCAA 744
                ||  |||||  |   ||  ||  ||   |  ||  |||||||   |     |   |  |||  ||   |  |
Db          709 ATATCGCAAGCACAAGTGTCAAGACTTGAAAAGGCTGCAATCAAACAGATGAATAAAAAT 768

Qy          745 ATTGA 749
                |||| |
Db          769 ATTCA 773
```

RESULT 5
ABK74144
ID ABK74144 standard; DNA; 381 BP.
XX
AC ABK74144;
XX
DT 13-AUG-2002 (first entry)
XX
DE Bacillus licheniformis genomic sequence tag (GST) #1435.
XX
KW Differential gene expression; genomic sequenced tag; GST;
KW altered culture condition; environmental stress;
KW physiological provocation; ds.
XX
OS Bacillus licheniformis.
XX
PN WO200229113-A2.
XX
PD 11-APR-2002.
XX
PF 05-OCT-2001; 2001WO-US031437.
XX
PR 06-OCT-2000; 2000US-00680598.
PR 27-MAR-2001; 2001US-0279526P.
XX
PA (NOVO) NOVOZYMES BIOTECH INC.
PA (NOVO) NOVOZYMES AS.
XX
PI Berka R, Clausen IG;
XX
DR WPI; 2002-416684/44.
XX
PT Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second Bacillus
PT cells, by using substrate containing Bacillus genomic sequenced tag
PT array.
XX
PS Claim 4; SEQ ID NO 1435; 200pp; English.
XX
CC The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive follow
CC -up characterisation is unnecessary, when one spot on an array equals one

```
CC   gene or one open reading frame, since sequence information is available.
CC   This sequence represents a genomic sequence tag (GST) used in the method
CC   of the invention. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 381 BP; 109 A; 91 C; 94 G; 87 T; 0 U; 0 Other;

  Query Match             23.0%;  Score 173.8;  DB 6;  Length 381;
  Best Local Similarity   66.9%;  Pred. No. 3.4e-45;
  Matches  247;  Conservative    0;  Mismatches  122;  Indels    0;  Gaps    0;

Qy          50 AAGTGAAAAGATTTGATAGCCAAGAGCCAAGCTGGCGATACGGATGCACGTGAGCTTCTCG 109
               |||||||||||  ||||     || |||||     |||||||     | ||| | || || |||
Db          13 AAGTGAAAAGAACTGATTAAAAACAGCCAGAACGGCGATCAAAAAGCAAGGGACCTCCTCA 72

Qy         110 TGAATAGCAATATCAGACTGGTCTGGTCCGTCGTCCAGCGCTTTATCAACCGCGGGTATG 169
               |  | |  || ||  | || || ||||| ||||| ||||| ||| | ||| | || ||||
Db          73 TAGAAAAAAACATGCGTCTTGTTTGGTCTGTCGTTCAGCGTTTTTTGAACAGAGGCTATG 132

Qy         170 AAGCGGATGATTTGTTTCAGATCGGTTGCATTGGCTTGCTCAAGGCCGTTGACAAGTTCG 229
               |  | || ||  | |||||  ||||| |||||  ||| |  | |||| || |||||  ||||
Db         133 AGCCTGACGACCTCTTTCAAATCGGCTGCATCGGCCTCTTGAAGTCGGTGGACAAATTCG 192

Qy         230 ATCTTTCGTACGATGTGAGATTTTCGACCTATGCGGTGCCAATGATCATCGGAGAAATTC 289
               |||||||  || ||  ||   | |||||  |||||  ||  ||  ||  |||||  ||||||  ||  ||||
Db         193 ATCTTTCCTATGACGTTCGGTTTTCCACCTACGCCGTTCCGATGATTATCGGCGAGATTC 252

Qy         290 AACGCTTTTTGCGCGATGACGGTACGGTTAAGGTCAGTCGATCGTTAAAAGAAACAGCGA 349
               | ||  |||  |   |  |||||||||  ||  ||  ||  ||  ||  |||  |  |||||||     |  |
Db         253 AGCGGTTTATCAGAGATGACGGAACCGTCAAAGTGAGCCGCTCGCTGAAAGAACTCGGCA 312

Qy         350 ATAAGGTGCGGCGATCAAAGGATGAATTGTACAAGCAATTCGGCCGTGCCCCCACGATCG 409
               | ||   | |||||   | ||  ||  ||   | |  ||||  |   |||||||     ||  ||  |
Db         313 ACAAAATCCGGCGGGCGAAAGACGAGCTTTCCAAGTCAAACGGCCGGATTCCGACCGTTC 372

Qy         410 CAGAAGTGG 418
                 |||  | |
Db         373 AGGAAATCG 381


RESULT 6
ABK74048
ID   ABK74048 standard; DNA; 631 BP.
XX
AC   ABK74048;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Bacillus licheniformis genomic sequence tag (GST) #1339.
XX
KW   Differential gene expression; genomic sequenced tag; GST;
KW   altered culture condition; environmental stress;
KW   physiological provocation; ds.
XX
OS   Bacillus licheniformis.
XX
PN   WO200229113-A2.
XX
PD   11-APR-2002.
XX
PF   05-OCT-2001; 2001WO-US031437.
XX
PR   06-OCT-2000; 2000US-00680598.
PR   27-MAR-2001; 2001US-0279526P.
XX
PA   (NOVO ) NOVOZYMES BIOTECH INC.
PA   (NOVO ) NOVOZYMES AS.
XX
```

PI Berka R, Clausen IG;
XX
DR WPI; 2002-416684/44.
XX
PT Monitoring differential expression of several genes in first Bacillus
PT cell relative to expression of same genes in one or more second Bacillus
PT cells, by using substrate containing Bacillus genomic sequenced tag
PT array.
XX
PS Claim 4; SEQ ID NO 1339; 200pp; English.
XX
CC The invention describes a method of monitoring differential expression of
CC genes in a first Bacillus cell relative to expression of the genes in
CC other Bacillus cells, comprising hybridising labelled nucleic acid probes
CC isolated from Bacillus cells to a substrate containing array of Bacillus
CC genomic sequenced tags (GST), examining the array, and determining
CC relative gene expression by an observed hybridisation reporter signal of
CC a spot in the array. The method is useful for measuring the expression of
CC genes in a first Bacillus cell relative to expression of the same genes
CC in one or more second Bacillus cells. The method is useful for monitoring
CC global expression of several genes from a Bacillus cell, discovering new
CC genes, identifying possible functions of unknown open reading frames and
CC monitoring gene copy number variation and stability. Monitoring changes
CC in expression of genes may be used to provide a representation of the way
CC in which Bacillus cells adapt to changes in culture conditions,
CC environmental stress or other physiological provocation. Extensive follow
CC -up characterisation is unnecessary, when one spot on an array equals one
CC gene or one open reading frame, since sequence information is available.
CC This sequence represents a genomic sequence tag (GST) used in the method
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 631 BP; 203 A; 121 C; 161 G; 146 T; 0 U; 0 Other;

Query Match 19.4%; Score 147; DB 6; Length 631;
Best Local Similarity 56.8%; Pred. No. 2.3e-36;
Matches 270; Conservative 0; Mismatches 205; Indels 0; Gaps 0;

```
Qy          77 AAGCTGGCGATACGGATGCACGTGAGCTTCTCGTGAATAGCAATATCAGACTGGTCTGGT 136
               |  |  ||  ||  ||      ||  |  ||     ||  ||  ||  ||||||  |  |  ||  ||||
Db          98 ATGAAGGAGACACAACAGCGAGAGAAAAGCTTGTTAACGGCAATTTGCGCCTTGTCTTAA 157

Qy         137 CCGTCGTCCAGCGCTTTATCAACCGCGGGTATGAAGCGGATGATTTGTTTCAGATCGGTT 196
                ||||  |  ||   |  ||||  ||||  |  ||   |   |  |  |||||  ||  ||  ||   ||||  |
Db         158 GCGTCATTCAAAGGTTTAACAACAGAGGAGAATATGTTGATGACTTATTCCAAGTCGGCT 217

Qy         197 GCATTGGCTTGCTCAAGGCCGTTGACAAGTTCGATCTTTCGTACGATGTGAGATTTTCGA 256
               ||||  ||   |   |  ||   |   ||||  ||  ||  ||  ||      ||  ||||  |   |||||  |
Db         218 GCATCGGACTAATGAAATCAATTGATAATTTTGACCTGAGCCACAATGTTAAGTTTTCAA 277

Qy         257 CCTATGCGGTGCCAATGATCATCGGAGAAATTCAACGCTTTTTGCGCGATGACGGTACGG 316
               |  ||||||  ||  ||||||||||||||||||||  |     |  |  ||||||||||  ||      ||
Db         278 CATATGCTGTACCAATGATCATCGGAGAAATCCGCAGATATTTGCGCGATAACAACCCGA 337

Qy         317 TTAAGGTCAGTCGATCGTTAAAAGAAACAGCGAATAAGGTGCGGCGATCAAAGGATGAAT 376
               |      |||    ||  ||   |  |   ||  |   |||  |  ||||  ||  ||  |    |   ||
Db         338 TCCGCGTCTCACGGTCACTCAGGGATATCGCGTACAAGGCGCTGCAAGTGAGAGAACGGC 397

Qy         377 TGTACAAGCAATTCGGCCGTGCCCCCACGATCGCAGAAGTGGCAGAAGCAGTGGGAATCA 436
               ||  ||    |      ||  |  |  ||  ||        ||   |  ||   |||    |  |  |  |
Db         398 TGATCAGTGAGACAAGCAGGGAGCCGACTGCTCAGGAGATCGCTAAAGAGCTTGAAGTGT 457

Qy         437 CGCCGGAGGAAGTAGTCTTTGCGCAAGAGGCAAGCAGAGCGCCTTCCTCCATCCATGAGA 496
               |  |  ||  |||  |  ||  |||||||   ||  |||  |      ||   |||    ||   |    ||||
Db         458 CCCATGAAGAAATCGTTTTTGCGCTTGACGCCATTCAAGATCCTGTATCTTTGTTTGAGC 517

Qy         497 CCGTTTTTGAAAATGACGGCGATCCCATCACACTGATCGATCAGATAGCGGATGA 551
                   |  |||    |   |  |   ||  |||||  ||      |  ||  |||||  ||    |||||
```

```
Db          518 CGATTTACAATGACGGAGGAGATCCGATTTATGTCATGGATCAAATCAGCGATGA 572


RESULT 7
ADH97013
ID   ADH97013 standard; DNA; 1110 BP.
XX
AC   ADH97013;
XX
DT   06-MAY-2004  (first entry)
XX
DE   S. pneumoniae RNA polymerase sigma-70 factor gene #2.
XX
KW   antibacterial; antiinflammatory; gastrointestinal; antiulcer;
KW   antidiarrhoeic; ophthalmological; enzyme inhibitor; antisense therapy;
KW   vaccine; microbial target; modulator; furuncle; pneumonia; gastritis;
KW   peptic ulcer disease; diarrhoea; meningitis; bacteraemia; conjunctivitis;
KW   toxic shock syndrome; gene; ds.
XX
OS   Streptococcus pneumoniae.
XX
PN   WO2003102190-A2.
XX
PD   11-DEC-2003.
XX
PF   02-JUN-2003; 2003WO-CA000786.
XX
PR   31-MAY-2002; 2002US-0384634P.
PR   31-MAY-2002; 2002US-0385157P.
PR   04-JUN-2002; 2002US-0385542P.
PR   04-JUN-2002; 2002US-0385611P.
PR   04-JUN-2002; 2002US-0385747P.
PR   04-JUN-2002; 2002US-0385750P.
PR   04-JUN-2002; 2002US-0385752P.
PR   04-JUN-2002; 2002US-0385773P.
PR   04-JUN-2002; 2002US-0385780P.
PR   04-JUN-2002; 2002US-0385785P.
PR   04-JUN-2002; 2002US-0385797P.
PR   05-JUN-2002; 2002US-0385962P.
PR   05-JUN-2002; 2002US-0386022P.
PR   05-JUN-2002; 2002US-0386024P.
PR   05-JUN-2002; 2002US-0386087P.
PR   05-JUN-2002; 2002US-0386141P.
PR   05-JUN-2002; 2002US-0386350P.
PR   05-JUN-2002; 2002US-0386586P.
PR   06-JUN-2002; 2002US-0386368P.
PR   06-JUN-2002; 2002US-0386369P.
PR   06-JUN-2002; 2002US-0386436P.
PR   06-JUN-2002; 2002US-0386441P.
PR   06-JUN-2002; 2002US-0386528P.
PR   06-JUN-2002; 2002US-0386573P.
PR   06-JUN-2002; 2002US-0386834P.
PR   31-JUL-2002; 2002US-0399839P.
PR   31-JUL-2002; 2002US-0399861P.
PR   31-JUL-2002; 2002US-0399969P.
PR   31-JUL-2002; 2002US-0399970P.
PR   31-JUL-2002; 2002US-0399983P.
PR   31-JUL-2002; 2002US-0399984P.
PR   31-JUL-2002; 2002US-0399985P.
PR   01-AUG-2002; 2002US-0400154P.
PR   01-AUG-2002; 2002US-0400230P.
PR   01-AUG-2002; 2002US-0400268P.
PR   01-AUG-2002; 2002US-0400363P.
PR   01-AUG-2002; 2002US-0400365P.
PR   01-AUG-2002; 2002US-0400374P.
PR   01-AUG-2002; 2002US-0400380P.
PR   01-AUG-2002; 2002US-0400433P.
PR   01-AUG-2002; 2002US-0400434P.
PR   01-AUG-2002; 2002US-0400436P.
```

PR 01-AUG-2002; 2002US-0400442P.
PR 01-AUG-2002; 2002US-0400463P.
XX
PA (AFFI-) AFFINIUM PHARM INC.
XX
PI Edwards A, Dharamsi A, Vedadi M, Vallee F, Awrey D, Beattie B;
PI Richards D, Domagala M, Mansoury K, Virag C, Buzadzija K;
PI Mcdonald M, Houston S, Arrowsmith C, Ouyang H, Nethery K, Ng I;
PI Kanagarajah D;
XX
DR WPI; 2004-071165/07.
DR P-PSDB; ADH97014.
XX
PT Compositions comprising recombinant polypeptide targets for pathogenic
PT bacteria, useful for designing modulators for preventing or treating a
PT disease or disorder associated with the species of origin for the
PT polypeptide.
XX
PS Claim 23; SEQ ID NO 204; 606pp; English.
XX
CC The invention relates to novel compositions (I) comprising isolated,
CC recombinant polypeptides, amino acid sequences having at least about 95%
CC identity with these or an amino acid sequence encoded by a polynucleotide
CC that hybridizes under stringent conditions to the complementary strand of
CC the polynucleotide encoding these polypeptides. The compositions and
CC polypeptides are useful as microbial targets for designing modulators for
CC the prevention or treatment of a disease or disorder associated with the
CC species of origin for the polypeptide, e.g. furuncle, pneumonia,
CC gastritis, peptic ulcer disease, diarrhoea, meningitis, bacteraemia,
CC conjunctivitis or toxic shock syndrome. The polypeptides are also useful
CC for diagnosing a patient suffering from a disease or disorder of a
CC pathogenic species, or for monitoring the effectiveness of an anti-
CC pathogenic treatment. This sequence corresponds to one of the DNA
CC sequences of the invention
XX
SQ Sequence 1110 BP; 341 A; 211 C; 267 G; 291 T; 0 U; 0 Other;

Query Match 10.3%; Score 78; DB 12; Length 1110;
Best Local Similarity 50.7%; Pred. No. 8.1e-14;
Matches 214; Conservative 0; Mismatches 205; Indels 3; Gaps 1;

```
Qy          31 CCATTTCTGACCAATGACCAAGTGAAAGATTTGATAGCCAAGAGCCAAGCTGGCGATACG 90
               ||  |  ||||||||||  | | ||||||  |||  |          |||||||  ||||
Db         328 CCTCTCTTGACCAATGAAGAGGAGAAAGAGTTGGCACTGGCTGTTGAAGCTGGTGATATC 387

Qy          91 GATGCACGTGAGCTTCTCGTGAATAGCAATATCAGACTGGTCTGGTCCGTCGTCCAGCGC 150
               ||  ||      | | ||| | | |  |||||  |   |  ||||     |||  | | | | |||
Db         388 GAAGCCAAACAACGTCTTGCGGAACGCAATCTTCGTTTGGTTGTTTCCATTGCCAAACGC 447

Qy         151 TTTATCAACCGCGGGTATGAAGCGGATGATTTGTTTCAGATCGGTTGCATTGGCTTGCTC 210
               | | ||   || ||       |     ||| ||| ||||    ||     || ||||||  |
Db         448 TATGTCGGTCGTGGCATGCAGTTCCTTGACTTGATTCAAGAAGGAAATATGGGCTTGATG 507

Qy         211 AAGGCCGTTGACAAGTTCGATCTTTCGTACGATGTGAGATTTTCGACCTATGCGGTGCCA 270
               |||||  ||||||||||||  ||   |||  | |   | |  |||||  ||  |||||
Db         508 AAGGCGGTTGACAAGTTTGACTATTCTAAAGGGTTCAAGTTTTCAACTTATGCAACTTGG 567

Qy         271 ATGATCATCGGAGAAAAT---TCAACGCTTTTTGCGCGATGACGGTACGGTTAAGGTCAGT 327
                 |||       |  ||   ||      ||  |  | | |   ||||  |      ||
Db         568 TGGATTCGTCAGGCTATCACTCGTGCTATTGCGGACCAAGCTCGTACCATCCGTATCCCA 627

Qy         328 CGATCGTTAAAAGAAACAGCGAATAAGGTGCGGCGATCAAAGGATGAATTGTACAAGCAA 387
                  |   |||||   |||||  ||  ||   | ||   |  |   |  ||
Db         628 GTTCACATGGTTGAAACTATCAATAAATTGGTTCGTGAACAGCGGAATCTCCTTCAAGAA 687

Qy         388 TTCGGCCGTGCCCCCACGATCGCAGAAGTGGCAGAAGCAGTGGGAATCACGCCGGAGGAA 447
               || ||  |  || ||    | | |  || |||  | |||  || ||  ||  ||   |
Db         688 TTGGGGCAAGATCCGACACCAGAACAGATTGCTGAACGAATGGATATGACACCTGATAAG 747
```

```
Qy          448 GT 449
                ||
Db          748 GT 749


RESULT 8
AEC13368
ID   AEC13368 standard; DNA; 1110 BP.
XX
AC   AEC13368;
XX
DT   20-OCT-2005  (first entry)
XX
DE   Streptococcus pneumoniae RNA polymerase sigma-70 factor gene.
XX
KW   protein purification; antibacterial; antimicrobial; infection;
KW   drug screening; RNA polymerase sigma-70 factor; gene; ss.
XX
OS   Streptococcus pneumoniae.
XX
PN   US2005181464-A1.
XX
PD   18-AUG-2005.
XX
PF   29-SEP-2004; 2004US-00953901.
XX
PR   04-APR-2002; 2002US-0369819P.
PR   04-APR-2002; 2002US-0369826P.
PR   04-APR-2002; 2002US-0369831P.
PR   04-APR-2002; 2002US-0370060P.
PR   08-APR-2002; 2002US-0370681P.
PR   08-APR-2002; 2002US-0370806P.
PR   08-APR-2002; 2002US-0370852P.
PR   08-APR-2002; 2002US-0370868P.
PR   09-APR-2002; 2002US-0370959P.
PR   09-APR-2002; 2002US-0370978P.
PR   09-APR-2002; 2002US-0371008P.
PR   09-APR-2002; 2002US-0371009P.
PR   09-APR-2002; 2002US-0371014P.
PR   09-APR-2002; 2002US-0371025P.
PR   09-APR-2002; 2002US-0371064P.
PR   09-APR-2002; 2002US-0371065P.
PR   09-APR-2002; 2002US-0371094P.
PR   09-APR-2002; 2002US-0371114P.
PR   09-APR-2002; 2002US-0371180P.
PR   09-APR-2002; 2002US-0371189P.
PR   31-MAY-2002; 2002US-0384634P.
PR   31-MAY-2002; 2002US-0385157P.
PR   04-JUN-2002; 2002US-0385542P.
PR   04-JUN-2002; 2002US-0385611P.
PR   04-JUN-2002; 2002US-0385747P.
PR   04-JUN-2002; 2002US-0385750P.
PR   04-JUN-2002; 2002US-0385752P.
PR   04-JUN-2002; 2002US-0385773P.
PR   04-JUN-2002; 2002US-0385780P.
PR   04-JUN-2002; 2002US-0385785P.
PR   04-JUN-2002; 2002US-0385797P.
PR   05-JUN-2002; 2002US-0385962P.
PR   05-JUN-2002; 2002US-0386022P.
PR   05-JUN-2002; 2002US-0386024P.
PR   05-JUN-2002; 2002US-0386087P.
PR   05-JUN-2002; 2002US-0386141P.
PR   05-JUN-2002; 2002US-0386350P.
PR   05-JUN-2002; 2002US-0386586P.
PR   06-JUN-2002; 2002US-0386368P.
PR   06-JUN-2002; 2002US-0386369P.
PR   06-JUN-2002; 2002US-0386436P.
PR   06-JUN-2002; 2002US-0386441P.
PR   06-JUN-2002; 2002US-0386528P.
```

PR 06-JUN-2002; 2002US-0386573P.
PR 06-JUN-2002; 2002US-0386834P.
PR 31-JUL-2002; 2002US-0399839P.
PR 31-JUL-2002; 2002US-0399861P.
PR 31-JUL-2002; 2002US-0399969P.
PR 31-JUL-2002; 2002US-0399970P.
PR 31-JUL-2002; 2002US-0399983P.
PR 31-JUL-2002; 2002US-0399984P.
PR 31-JUL-2002; 2002US-0399985P.
PR 01-AUG-2002; 2002US-0400154P.
PR 01-AUG-2002; 2002US-0400230P.
PR 01-AUG-2002; 2002US-0400268P.
PR 01-AUG-2002; 2002US-0400363P.
PR 01-AUG-2002; 2002US-0400365P.
PR 01-AUG-2002; 2002US-0400374P.
PR 01-AUG-2002; 2002US-0400380P.
PR 01-AUG-2002; 2002US-0400433P.
PR 01-AUG-2002; 2002US-0400434P.
PR 01-AUG-2002; 2002US-0400436P.
PR 01-AUG-2002; 2002US-0400442P.
PR 01-AUG-2002; 2002US-0400463P.
PR 04-APR-2003; 2003WO-CA000465.
PR 08-APR-2003; 2003WO-CA000482.
PR 08-APR-2003; 2003WO-CA000483.
PR 02-JUN-2003; 2003WO-CA000786.
XX
PA (AFFI-) AFFINIUM PHARM INC.
XX
PI Edwards A, Dharamsi A, Vedadi M, Alam MZ, Arrowsmith C, Awrey DE;
PI Beattie B, Buzadzija K, Clarke T, Domagala M, Houston S;
PI Kanagarajah D, Li Q, Mansoury K, Mcdonald M, Nethery-Brokx K, Ng I;
PI Ouyang H, Richards D, Vallee F, Virag C;
XX
DR WPI; 2005-628190/64.
DR P-PSDB; AEC13369.
XX
PT Novel crystallized, recombinant bacterial polypeptide, useful as targets
PT for pathogenic bacteria such as Helicobacter pylori, Staphylococcus
PT aureus, for detecting pathogenic species in biological sample, and in
PT drug designing.
XX
PS Claim 85; SEQ ID NO 204; 637pp; English.
XX
CC The invention relates to a composition (I) comprising purified
CC polypeptides from bacteria. Also described: (1) a crystallized,
CC recombinant polypeptide comprising an amino acid sequence of (I), where
CC the polypeptide is in crystal form; (2) a crystallized complex comprising
CC the crystallized, recombinant polypeptide and a co-factor or a small
CC organic molecule, where the complex is in crystal form; and (3) a host
CC cell comprising a nucleic acid encoding a polypeptide of (I), where a
CC culture of the host cell produces at least about 1 mg of the polypeptide
CC per liter of culture and the polypeptide is at least about one-third
CC soluble as measured by gel electrophoresis. (I) can be used as a target
CC for pathogenic bacteria, useful for detecting the presence of a
CC pathogenic species in a biological sample. (I) is useful for monitoring
CC the effectiveness of anti-pathogenic treatments in an individual
CC suffering from a disease or disorder caused by a pathogenic bacteria,
CC such as infections. (I) is also useful in drug design and screening, for
CC identifying inhibitors of (I), for designing a potential compound that is
CC useful for treating or preventing pathogenic diseases or disorders, for
CC assessing the activity of small molecules and other modulators in in
CC vitro assay, and for developing antimicrobial agents. The present
CC sequence represents a Streptococcus pneumoniae RNA polymerase sigma-70
CC factor gene, which is used in an example from the present invention.
XX
SQ Sequence 1110 BP; 341 A; 211 C; 267 G; 291 T; 0 U; 0 Other;

Query Match 10.3%; Score 78; DB 18; Length 1110;
Best Local Similarity 50.7%; Pred. No. 8.1e-14;

```
  Matches  214;  Conservative    0;  Mismatches  205;  Indels    3;  Gaps    1;

Qy          31 CCATTTCTGACCAATGACCAAGTGAAAGATTTGATAGCCAAGAGCCAAGCTGGCGATACG 90
               ||  |  ||||||||||  | | ||||||| ||| |          |||||||_||||
Db         328 CCTCTCTTGACCAATGAAGAGGAGAAAGAGTTGGCACTGGCTGTTGAAGCTGGTGATATC 387

Qy          91 GATGCACGTGAGCTTCTCGTGAATAGCAATATCAGACTGGTCTGGTCCGTCGTCCAGCGC 150
               ||  ||      | | ||| | | |  |||||  |   |  ||||     ||| | | | | |||
Db         388 GAAGCCAAACAACGTCTTGCGGAACGCAATCTTCGTTTGGTTGTTTCCATTGCCAAACGC 447

Qy         151 TTTATCAACCGCGGGTATGAAGCGGATGATTTGTTTCAGATCGGTTGCATTGGCTTGCTC 210
               | | ||   || ||      |       ||| ||| ||||     ||     || ||||||  |
Db         448 TATGTCGGTCGTGGCATGCAGTTCCTTGACTTGATTCAAGAAGGAAATATGGGCTTGATG 507

Qy         211 AAGGCCGTTGACAAGTTCGATCTTTCGTACGATGTGAGATTTTCGACCTATGCGGTGCCA 270
               ||||| ||||||||||||| ||   ||| | | |  | |  |||||  || |||||
Db         508 AAGGCGGTTGACAAGTTTGACTATTCTAAAGGGTTCAAGTTTTCAACTTATGCAACTTGG 567

Qy         271 ATGATCATCGGAGAAAT---TCAACGCTTTTTGCGCGATGACGGTACGGTTAAGGTCAGT 327
                 |||         |  ||  ||       ||  |  | | |    ||||  |      ||
Db         568 TGGATTCGTCAGGCTATCACTCGTGCTATTGCGGACCAAGCTCGTACCATCCGTATCCCA 627

Qy         328 CGATCGTTAAAAGAAACAGCGAATAAGGTGCGGCGATCAAAGGATGAATTGTACAAGCAA 387
                  |    |||||   |||||  ||   ||   | ||    |  |    |  ||
Db         628 GTTCACATGGTTGAAACTATCAATAAATTGGTTCGTGAACAGCGGAATCTCCTTCAAGAA 687

Qy         388 TTCGGCCGTGCCCCCACGATCGCAGAAGTGGCAGAAGCAGTGGGAATCACGCCGGAGGAA 447
               || || |  |  || ||    | | |  | || ||| | |||  || || || ||  |
Db         688 TTGGGGCAAGATCCGACACCAGAACAGATTGCTGAACGAATGGATATGACACCTGATAAG 747

Qy         448 GT 449
               ||
Db         748 GT 749
```

```
RESULT 9
ADK44581
ID   ADK44581 standard; DNA; 915 BP.
XX
AC   ADK44581;
XX
DT   24-FEB-2005 (first entry)
XX
DE   Streptococcus pneumoniae gene, Seq ID No 1096.
XX
KW   ds; gene; Antibacterial; Gene therapy; Vaccine; Streptococcus pneumoniae.
XX
OS   Streptococcus pneumoniae.
XX
PN   US6699703-B1.
XX
PD   02-MAR-2004.
XX
PF   26-MAY-2000; 2000US-00583110.
XX
PR   02-JUL-1997;    97US-0051553P.
PR   12-MAY-1998;    98US-0085131P.
PR   30-JUN-1998;    98US-00107433.
XX
PA   (GENO-) GENOME THERAPEUTICS CORP.
XX
PI   Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;
XX
DR   WPI; 2004-212399/20.
DR   P-PSDB; ADK47242.
XX
PT   New nucleic acid molecules and polypeptides useful for diagnosing,
PT   preventing and treating pathological conditions resulting from bacterial
PT   infection, e.g. Streptococcus pneumoniae infection, and in drug
```

```
PT   screening.
XX
PS   Disclosure; SEQ ID NO 1096; 301pp; English.
XX
CC   The invention relates to isolated Streptococcus pneumoniae nucleic acids
CC   and polypeptides. The nucleic acids and proteins are useful for
CC   diagnosing, preventing and treating pathological conditions resulting
CC   from bacterial infection, such as S. pneumoniae infection. These may also
CC   be used for drug screening procedures. The present sequence represents a
CC   Streptococcus pneumoniae nucleic acid of the invention. Note: The
CC   sequence data for this patent did not appear in the printed specification
CC   but was obtained in electronic format directly from USPTO at
CC   seqdata.uspto.gov/sequence.html.
XX
SQ   Sequence 915 BP; 279 A; 175 C; 220 G; 241 T; 0 U; 0 Other;

  Query Match             10.1%;  Score 76.4;  DB 12;  Length 915;
  Best Local Similarity   50.5%;  Pred. No. 2.4e-13;
  Matches  213;  Conservative    0;  Mismatches  206;  Indels    3;  Gaps    1;

Qy          31 CCATTTCTGACCAATGACCAAGTGAAAGATTTGATAGCCAAGAGCCAAGCTGGCGATACG 90
               ||  |  ||||||||||  | | |||||| ||| |||||||||||||| |||| |
Db         328 CCTCTCTTGACCAATGAAGAGGAGAAAGAGTTGGCACTGGCTGTTGAAGCTGGTGATATC 387

Qy          91 GATGCACGTGAGCTTCTCGTGAATAGCAATATCAGACTGGTCTGGTCCGTCGTCCAGCGC 150
               || ||      | | ||| | | |    |||| |  | |||      ||| | | | | |||
Db         388 GAAGCCAAACAACGTCTTGCGGAAGCCAATCTTCGTTTGGTTGTTTCCATTGCCAAACGC 447

Qy         151 TTTATCAACCGCGGGTATGAAGCGGATGATTTGTTTCAGATCGGTTGCATTGGCTTGCTC 210
               | | ||   || ||     |      ||| ||| ||||     ||     || |||||| |
Db         448 TATGTCGGTCGTGGTATGCAGTTCCTTGACTTGATTCAAGAAGGAAATATGGGCTTGATG 507

Qy         211 AAGGCCGTTGACAAGTTCGATCTTTCGTACGATGTGAGATTTTCGACCTATGCGGTGCCA 270
               ||||| ||||||||||||  ||    ||| | |    | |  ||||| || |||||
Db         508 AAGGCGGTTGACAAGTTTGACTATTCTAAAGGGTTCAAGTTTTCAACTTATGCAACTTGG 567

Qy         271 ATGATCATCGGAGAAATTCAACGCTTTTTGCGCGATGACG---GTACGGTTAAGGTCAGT 327
                 |||      |  ||     ||   | |    ||  | |   ||||  |     ||
Db         568 TGGATTCGTCAGGCTATCACTCGTGCTATTGCAGACCAAGCTCGTACCATCCGTATCCCA 627

Qy         328 CGATCGTTAAAAGAAACAGCGAATAAGGTGCGGCGATCAAAGGATGAATTGTACAAGCAA 387
                   |    |||||    |||||  ||   ||   | ||     |  |     |  ||
Db         628 GTTCACATGGTTGAAACTATCAATAAATTGGTTCGTGAACAGCGGAATCTCCTTCAAGAA 687

Qy         388 TTCGGCCGTGCCCCCACGATCGCAGAAGTGGCAGAAGCAGTGGGAATCACGCCGGAGGAA 447
               || || |  |  || ||     | | |  | || |||  | |||  || || || ||  |
Db         688 TTGGGGCAAGATCCGACACCAGAACAGATTGCTGAACGAATGGATATGACACCTGATAAG 747

Qy         448 GT 449
               ||
Db         748 GT 749


RESULT 10
AEJ68509
ID   AEJ68509 standard; DNA; 915 BP.
XX
AC   AEJ68509;
XX
DT   05-OCT-2006  (first entry)
XX
DE   Streptococcus pneumoniae strain 14453 protein-encoding DNA, SEQ:1096.
XX
KW   Vaccine; diagnosis; drug discovery; protein production;
KW   bacterial infection; Streptococcus pneumoniae infection;
KW   bacterial meningitis; bacterial pneumonia; bacteremia; otitis media;
KW   antibacterial; neuroprotective; antiinflammatory; respiratory-gen.;
KW   auditory; gene; ds.
XX
```

OS Streptococcus pneumoniae; strain 14453.
XX
FH Key Location/Qualifiers
FT CDS 1..915
FT /*tag= a
FT /product= "Streptococcus pneumoniae protein SEQ ID NO:3757"
XX
PN US7074914-B1.
XX
PD 11-JUL-2006.
XX
PF 30-DEC-2004; 2004US-00028099.
XX
PR 02-JUL-1997; 97US-0051553P.
PR 12-MAY-1998; 98US-0085131P.
PR 30-JUN-1998; 98US-00107433.
PR 26-MAY-2000; 2000US-00583110.
PR 14-AUG-2003; 2003US-00640833.
XX
PA (SNFI) SANOFI PASTEUR LTEE.
XX
PI Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;
XX
DR WPI; 2006-500481/51.
DR P-PSDB; AEJ71170.
XX
PT New isolated nucleic acid and polypeptide from Streptococcus pneumoniae,
PT useful for diagnosing, preventing, or treating pathological conditions
PT resulting from bacterial infections, e.g. S. pneumoniae infection.
XX
PS Example; SEQ ID NO 1096; 29pp; English.
XX
CC The invention relates to an isolated nucleic acid, especially (AEJ68056),
CC which encodes the Streptococcus pneumoniae protein of AEJ70717. This
CC nucleic acid is one of 2661 disclosed protein-encoding nucleic acids
CC (AEJ67414-AEJ70074) isolated from a Streptococcus pneumoniae strain 14453
CC genomic library whose predicted products (AEJ70075-AEJ72735) exhibit
CC homology to known prokaryotic, eukaryotic or archaeal open reading frames
CC (ORFs) or proteins. The invention also relates to a recombinant
CC expression vector comprising the nucleic acid of the invention operably
CC linked to a transcription regulatory element; and a host cell comprising
CC the recombinant expression vector. The Streptococcus pneumoniae nucleic
CC acids and proteins of the invention are useful for diagnosing,
CC preventing, or treating pathological conditions resulting from bacterial
CC infections, especially infections caused by Streptococcus pneumoniae such
CC as meningitis, bacteremia, pneumonia and otitis media. They may also be
CC used in vaccine compositions for the treatment of Streptococcus
CC pneumoniae infections and as targets for antibacterial drugs.
CC Additionally the nucleic acids are useful in the production of
CC commercially important proteins such as enzymes for use in fermentation
CC reactions or in the production of commercially useful metabolites. The
CC present sequence represents a Streptococcus pneumoniae strain 14453
CC protein-encoding nucleic acid which was identified in the exemplification
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=7074914B1.
XX
SQ Sequence 915 BP; 279 A; 175 C; 220 G; 241 T; 0 U; 0 Other;

Query Match 10.1%; Score 76.4; DB 21; Length 915;
Best Local Similarity 50.5%; Pred. No. 2.4e-13;
Matches 213; Conservative 0; Mismatches 206; Indels 3; Gaps 1;

```
Qy          31 CCATTTCTGACCAATGACCAAGTGAAAGATTTGATAGCCAAGAGCCAAGCTGGCGATACG 90
               ||  |  ||||||||||  | | ||||||| |||  |           |||||||  ||||
Db         328 CCTCTCTTGACCAATGAAGAGGAGAAAGAGTTGGCACTGGCTGTTGAAGCTGGTGATATC 387
```

```
Qy          91 GATGCACGTGAGCTTCTCGTGAATAGCAATATCAGACTGGTCTGGTCCGTCGTCCAGCGC 150
               || ||       | | ||| | | |    |||| |   |  ||||     ||| | | | | |||
Db         388 GAAGCCAAACAACGTCTTGCGGAAGCCAATCTTCGTTTGGTTGTTTCCATTGCCAAACGC 447

Qy         151 TTTATCAACCGCGGGTATGAAGCGGATGATTTGTTTCAGATCGGTTGCATTGGCTTGCTC 210
               | | ||    || ||       |      ||| ||| ||||     ||     || ||||||  |
Db         448 TATGTCGGTCGTGGTATGCAGTTCCTTGACTTGATTCAAGAAGGAAATATGGGCTTGATG 507

Qy         211 AAGGCCGTTGACAAGTTCGATCTTTCGTACGATGTGAGATTTTCGACCTATGCGGTGCCA 270
               ||||| ||||||||||||  ||    |||  | |    | |  |||||  ||  |||||
Db         508 AAGGCGGTTGACAAGTTTGACTATTCTAAAGGGTTCAAGTTTTCAACTTATGCAACTTGG 567

Qy         271 ATGATCATCGGAGAAATTCAACGCTTTTTGCGCGATGACG---GTACGGTTAAGGTCAGT 327
                 |||        |  ||     ||   | |    ||  | |    ||||  |      ||
Db         568 TGGATTCGTCAGGCTATCACTCGTGCTATTGCAGACCAAGCTCGTACCATCCGTATCCCA 627

Qy         328 CGATCGTTAAAAGAAACAGCGAATAAGGTGCGGCGATCAAAGGATGAATTGTACAAGCAA 387
                   |    |||||    |||||  ||   ||    | ||      |   |    |  ||
Db         628 GTTCACATGGTTGAAACTATCAATAAATTGGTTCGTGAACAGCGGAATCTCCTTCAAGAA 687

Qy         388 TTCGGCCGTGCCCCCACGATCGCAGAAGTGGCAGAAGCAGTGGGAATCACGCCGGAGGAA 447
               || || |  |  || ||     | | |  | || ||| | |||  || || || ||  |
Db         688 TTGGGGCAAGATCCGACACCAGAACAGATTGCTGAACGAATGGATATGACACCTGATAAG 747

Qy         448 GT 449
               ||
Db         748 GT 749
```

```
RESULT 11
AEJ75484
ID   AEJ75484 standard; DNA; 915 BP.
XX
AC   AEJ75484;
XX
DT   05-OCT-2006 (first entry)
XX
DE   Streptococcus pneumoniae strain 14453 protein-encoding DNA, SEQ:1096.
XX
KW   Vaccine; diagnosis; drug discovery; protein production;
KW   bacterial infection; Streptococcus pneumoniae infection;
KW   bacterial meningitis; bacterial pneumonia; bacteremia; otitis media;
KW   antibacterial; neuroprotective; antiinflammatory; respiratory-gen.;
KW   auditory; gene; ds.
XX
OS   Streptococcus pneumoniae; strain 14453.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..915
FT                   /*tag= a
FT                   /product= "Streptococcus pneumoniae protein SEQ ID
FT                   NO:3757"
XX
PN   US7081530-B1.
XX
PD   25-JUL-2006.
XX
PF   30-DEC-2004; 2004US-00028291.
XX
PR   02-JUL-1997;    97US-0051553P.
PR   12-MAY-1998;    98US-0085131P.
PR   30-JUN-1998;    98US-00107433.
PR   26-MAY-2000; 2000US-00583110.
PR   14-AUG-2003; 2003US-00640833.
XX
PA   (SNFI ) SANOFI PASTEUR LTEE.
XX
PI   Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;
XX
```

DR WPI; 2006-518920/53.
DR P-PSDB; AEJ78145.
XX
PT New isolated Streptococcus pneumoniae nucleic acid, useful as a molecular
PT target for detecting, diagnosing, preventing, or treating a pathological
PT condition resulting from bacterial infection.
XX
PS Example; SEQ ID NO 1096; 29pp; English.
XX
CC The invention relates to an isolated nucleic acid, especially (AEJ75443),
CC which encodes the Streptococcus pneumoniae protein of AEJ78104. This
CC nucleic acid is one of 2661 disclosed protein-encoding nucleic acids
CC (AEJ74389-AEJ77049) isolated from a Streptococcus pneumoniae strain 14453
CC genomic library whose predicted products (AEJ77050-AEJ79710) exhibit
CC homology to known prokaryotic, eukaryotic or archaeal open reading frames
CC (ORFs) or proteins. The invention also relates to a recombinant
CC expression vector comprising the nucleic acid of the invention operably
CC linked to a transcription regulatory element; and a host cell comprising
CC the recombinant expression vector. The Streptococcus pneumoniae nucleic
CC acids and proteins of the invention are useful for diagnosing,
CC preventing, or treating pathological conditions resulting from bacterial
CC infections, especially infections caused by Streptococcus pneumoniae such
CC as meningitis, bacteremia, pneumonia and otitis media. They may also be
CC used in vaccine compositions for the treatment of Streptococcus
CC pneumoniae infections and as targets for antibacterial drugs.
CC Additionally the nucleic acids are useful in the production of
CC commercially important proteins such as enzymes for use in fermentation
CC reactions or in the production of commercially useful metabolites. The
CC present sequence represents a Streptococcus pneumoniae strain 14453
CC protein-encoding nucleic acid which was identified in the exemplification
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=7081530B1.
XX
SQ Sequence 915 BP; 279 A; 175 C; 220 G; 241 T; 0 U; 0 Other;

Query Match 10.1%; Score 76.4; DB 21; Length 915;
Best Local Similarity 50.5%; Pred. No. 2.4e-13;
Matches 213; Conservative 0; Mismatches 206; Indels 3; Gaps 1;

```
Qy          31 CCATTTCTGACCAATGACCAAGTGAAAGATTTGATAGCCAAGAGCCAAGCTGGCGATACG 90
               ||  |  ||||||||||  | | ||||||| ||| |          |||||||||||
Db         328 CCTCTCTTGACCAATGAAGAGGAGAAAGAGTTGGCACTGGCTGTTGAAGCTGGTGATATC 387

Qy          91 GATGCACGTGAGCTTCTCGTGAATAGCAATATCAGACTGGTCTGGTCCGTCGTCCAGCGC 150
               ||  ||      | | ||| | | |    |||| |   |  ||||     ||| | | | | |||
Db         388 GAAGCCAAACAACGTCTTGCGGAAGCCAATCTTCGTTTGGTTGTTTCCATTGCCAAACGC 447

Qy         151 TTTATCAACCGCGGGTATGAAGCGGATGATTTGTTTCAGATCGGTTGCATTGGCTTGCTC 210
               | | ||   || ||       |      ||| ||| ||||    ||     || |||||| |
Db         448 TATGTCGGTCGTGGTATGCAGTTCCTTGACTTGATTCAAGAAGGAAATATGGGCTTGATG 507

Qy         211 AAGGCCGTTGACAAGTTCGATCTTTCGTACGATGTGAGATTTTCGACCTATGCGGTGCCA 270
               ||||| ||||||||||||| ||   |||  | |   | |  |||||  || |||||
Db         508 AAGGCGGTTGACAAGTTTGACTATTCTAAAGGGTTCAAGTTTTCAACTTATGCAACTTGG 567

Qy         271 ATGATCATCGGAGAAATTCAACGCTTTTTGCGCGATGACG---GTACGGTTAAGGTCAGT 327
                 |||       |  ||    ||   | |    ||  | |   ||||  |      ||
Db         568 TGGATTCGTCAGGCTATCACTCGTGCTATTGCAGACCAAGCTCGTACCATCCGTATCCCA 627

Qy         328 CGATCGTTAAAAGAAACAGCGAATAAGGTGCGGCGATCAAAGGATGAATTGTACAAGCAA 387
                   |    |||||   |||||  ||   ||  | ||    |  |    |  ||
Db         628 GTTCACATGGTTGAAACTATCAATAAATTGGTTCGTGAACAGCGGAATCTCCTTCAAGAA 687

Qy         388 TTCGGCCGTGCCCCCACGATCGCAGAAGTGGCAGAAGCAGTGGGAATCACGCCGGAGGAA 447
               || ||| |  | || ||    | | |  || ||| |  ||| || ||| ||  |
Db         688 TTGGGGCAAGATCCGACACCAGAACAGATTGCTGAACGAATGGATATGACACCTGATAAG 747
```

```
Qy          448 GT 449
                ||
Db          748 GT 749


RESULT 12
AEJ82844
ID   AEJ82844 standard; DNA; 915 BP.
XX
AC   AEJ82844;
XX
DT   19-OCT-2006 (first entry)
XX
DE   Streptococcus pneumoniae strain 14453 protein-encoding DNA, SEQ:1096.
XX
KW   Vaccine; diagnosis; drug discovery; protein production;
KW   bacterial infection; Streptococcus pneumoniae infection;
KW   bacterial meningitis; bacterial pneumonia; bacteremia; otitis media;
KW   antibacterial; neuroprotective; antiinflammatory; respiratory-gen.;
KW   auditory; gene; ds.
XX
OS   Streptococcus pneumoniae; strain 14453.
XX
FH   Key                 Location/Qualifiers
FT   CDS                 1..915
FT                       /*tag= a
FT                       /product= "Streptococcus pneumoniae protein SEQ ID
FT                       NO:3757"
XX
PN   US7098023-B1.
XX
PD   29-AUG-2006.
XX
PF   30-DEC-2004; 2004US-00027878.
XX
PR   02-JUL-1997;    97US-0051553P.
PR   12-MAY-1998;    98US-0085131P.
PR   30-JUN-1998;    98US-00107433.
PR   26-MAY-2000; 2000US-00583110.
PR   14-AUG-2003; 2003US-00640833.
XX
PA   (SNFI ) SANOFI PASTEUR LTD.
XX
PI   Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;
XX
DR   WPI; 2006-584390/60.
DR   P-PSDB; AEJ85505.
XX
PT   New isolated Streptococcus pneumoniae nucleic acid and polypeptide,
PT   useful as vaccines and as targets for diagnosing, preventing, or treating
PT   pathological conditions resulting from S. pneumoniae bacterial infection.
XX
PS   Example; SEQ ID NO 1096; 29pp; English.
XX
CC   The invention relates to an isolated nucleic acid, especially (AEJ84330),
CC   which encodes the Streptococcus pneumoniae protein of AEJ86991. This
CC   nucleic acid is one of 2661 disclosed protein-encoding nucleic acids
CC   (AEJ81749-AEJ84409) isolated from a Streptococcus pneumoniae strain 14453
CC   genomic library whose predicted products (AEJ84410-AEJ87070) exhibit
CC   homology to known prokaryotic, eukaryotic or archaeal open reading frames
CC   (ORFs) or proteins. The invention also relates to a recombinant
CC   expression vector comprising the nucleic acid of the invention operably
CC   linked to a transcription regulatory element; and a host cell comprising
CC   the recombinant expression vector. The Streptococcus pneumoniae nucleic
CC   acids and proteins of the invention are useful for diagnosing,
CC   preventing, or treating pathological conditions resulting from bacterial
CC   infections, especially infections caused by Streptococcus pneumoniae such
CC   as meningitis, bacteremia, pneumonia and otitis media. They may also be
CC   used in vaccine compositions for the treatment of Streptococcus
```

```
CC   pneumoniae infections and as targets for antibacterial drugs.
CC   Additionally the nucleic acids are useful in the production of
CC   commercially important proteins such as enzymes for use in fermentation
CC   reactions or in the production of commercially useful metabolites. The
CC   present sequence represents a Streptococcus pneumoniae strain 14453
CC   protein-encoding nucleic acid which was identified in the exemplification
CC   of the invention. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from the US patent office at
CC   seqdata.uspto.gov/sequence.html?DocID=7098023B1.
XX
SQ   Sequence 915 BP; 279 A; 175 C; 220 G; 241 T; 0 U; 0 Other;

  Query Match             10.1%;  Score 76.4;  DB 21;  Length 915;
  Best Local Similarity   50.5%;  Pred. No. 2.4e-13;
  Matches  213;  Conservative    0;  Mismatches  206;  Indels    3;  Gaps    1;

Qy          31 CCATTTCTGACCAATGACCAAGTGAAAGATTTGATAGCCAAGAGCCAAGCTGGCGATACG 90
               ||   | ||||||||||   | | |||||| ||| | |||||||| ||||
Db         328 CCTCTCTTGACCAATGAAGAGGAGAAAGAGTTGGCACTGGCTGTTGAAGCTGGTGATATC 387

Qy          91 GATGCACGTGAGCTTCTCGTGAATAGCAATATCAGACTGGTCTGGTCCGTCGTCCAGCGC 150
               || ||      | | ||| | | |    |||| | | |||| ||| | | | | |||
Db         388 GAAGCCAAACAACGTCTTGCGGAAGCCAATCTTCGTTTGGTTGTTTCCATTGCCAAACGC 447

Qy         151 TTTATCAACCGCGGGTATGAAGCGGATGATTTGTTTCAGATCGGTTGCATTGGCTTGCTC 210
               | | ||   || ||      |      ||| ||| ||||    ||     || |||||| |
Db         448 TATGTCGGTCGTGGTATGCAGTTCCTTGACTTGATTCAAGAAGGAAATATGGGCTTGATG 507

Qy         211 AAGGCCGTTGACAAGTTCGATCTTTCGTACGATGTGAGATTTTCGACCTATGCGGTGCCA 270
               ||||| ||||||||||||| ||    ||| | | |  | |  |||||| || ||||||
Db         508 AAGGCGGTTGACAAGTTTGACTATTCTAAAGGGTTCAAGTTTTCAACTTATGCAACTTGG 567

Qy         271 ATGATCATCGGAGAAATTCAACGCTTTTTGCGCGATGACG---GTACGGTTAAGGTCAGT 327
                 |||       |  ||    ||   | |    ||  | |   ||||  |      ||
Db         568 TGGATTCGTCAGGCTATCACTCGTGCTATTGCAGACCAAGCTCGTACCATCCGTATCCCA 627

Qy         328 CGATCGTTAAAAGAAACAGCGAATAAGGTGCGGCGATCAAAGGATGAATTGTACAAGCAA 387
                   |    |||||   |||||  ||  ||  | ||  |  |     |  ||
Db         628 GTTCACATGGTTGAAACTATCAATAAATTGGTTCGTGAACAGCGGAATCTCCTTCAAGAA 687

Qy         388 TTCGGCCGTGCCCCCACGATCGCAGAAGTGGCAGAAGCAGTGGGAATCACGCCGGAGGAA 447
               || || |  |  || ||     | | |  | || ||| | ||| || || || ||  |
Db         688 TTGGGGCAAGATCCGACACCAGAACAGATTGCTGAACGAATGGATATGACACCTGATAAG 747

Qy         448 GT 449
               ||
Db         748 GT 749


RESULT 13
AEL05163
ID   AEL05163 standard; DNA; 915 BP.
XX
AC   AEL05163;
XX
DT   30-NOV-2006  (first entry)
XX
DE   Streptococcus pneumoniae strain 14453 protein-encoding DNA, SEQ:1096.
XX
KW   Vaccine; diagnosis; drug discovery; protein production;
KW   bacterial infection; Streptococcus pneumoniae infection;
KW   bacterial meningitis; bacterial pneumonia; bacteremia; otitis media;
KW   antibacterial; neuroprotective; antiinflammatory; respiratory-gen.;
KW   auditory; gene; ds.
XX
OS   Streptococcus pneumoniae; strain 14453.
XX
FH   Key             Location/Qualifiers
```

```
FT   CDS                 1. .915
FT                       /*tag= a
FT                       /product= "Streptococcus pneumoniae protein SEQ ID
FT                       NO:3757"
XX
PN   US7115731-B1.
XX
PD   03-OCT-2006.
XX
PF   30-DEC-2004; 2004US-00027399.
XX
PR   02-JUL-1997;    97US-0051553P.
PR   12-MAY-1998;    98US-0085131P.
PR   30-JUN-1998;    98US-00107433.
PR   26-MAY-2000; 2000US-00583110.
PR   14-AUG-2003; 2003US-00640833.
XX
PA   (SNFI ) SANOFI PASTEUR LTD.
XX
PI   Doucette-Stamm L, Bush D, Zeng O, Opperman T, Houseweart CE;
XX
DR   WPI; 2006-744050/76.
DR   P-PSDB; AEL07824.
XX
PT   New nucleic acid encoding Streptococcus pneumoniae polypeptide, useful
PT   for detecting, preventing, and treating pathological conditions resulting
PT   from bacterial infection.
XX
PS   Example; SEQ ID NO 1096; 29pp; English.
XX
CC   The invention relates to an isolated nucleic acid, especially (AEL05123),
CC   which encodes the Streptococcus pneumoniae protein of AEL07784. This
CC   nucleic acid is one of 2661 disclosed protein-encoding nucleic acids
CC   (AEL04068-AEL06728) isolated from a Streptococcus pneumoniae strain 14453
CC   genomic library whose predicted products (AEL06729-AEL09389) exhibit
CC   homology to known prokaryotic, eukaryotic or archaeal open reading frames
CC   (ORFs) or proteins. The invention also relates to a recombinant
CC   expression vector comprising the nucleic acid of the invention operably
CC   linked to a transcription regulatory element; and a host cell comprising
CC   the recombinant expression vector. The Streptococcus pneumoniae nucleic
CC   acids and proteins of the invention are useful for diagnosing,
CC   preventing, or treating pathological conditions resulting from bacterial
CC   infections, especially infections caused by Streptococcus pneumoniae such
CC   as meningitis, bacteremia, pneumonia and otitis media. They may also be
CC   used in vaccine compositions for the treatment of Streptococcus
CC   pneumoniae infections and as targets for antibacterial drugs.
CC   Additionally the nucleic acids are useful in the production of
CC   commercially important proteins such as enzymes for use in fermentation
CC   reactions or in the production of commercially useful metabolites. The
CC   present sequence represents a Streptococcus pneumoniae strain 14453
CC   protein-encoding nucleic acid which was identified in the exemplification
CC   of the invention. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from the US patent office at
CC   seqdata.uspto.gov/sequence.html?DocID=7115731B1.
XX
SQ   Sequence 915 BP; 279 A; 175 C; 220 G; 241 T; 0 U; 0 Other;

  Query Match             10.1%;  Score 76.4;  DB 21;  Length 915;
  Best Local Similarity   50.5%;  Pred. No. 2.4e-13;
  Matches  213;  Conservative    0;  Mismatches  206;  Indels    3;  Gaps    1;

Qy          31 CCATTTCTGACCAATGACCAAGTGAAAGATTTGATAGCCAAGAGCCAAGCTGGCGATACG 90
               ||  |  ||||||||||  | | ||||||| |||  |          |||||||_||||
Db         328 CCTCTCTTGACCAATGAAGAGGAGAAAGAGTTGGCACTGGCTGTTGAAGCTGGTGATATC 387

Qy          91 GATGCACGTGAGCTTCTCGTGAATAGCAATATCAGACTGGTCTGGTCCGTCGTCCAGCGC 150
               || ||      | | ||| | | |    |||| |  |  ||||     ||| | | | | |||
Db         388 GAAGCCAAACAACGTCTTGCGGAAGCCAATCTTCGTTTGGTTGTTTCCATTGCCAAACGC 447
```

```
Qy          151 TTTATCAACCGCGGGTATGAAGCGGATGATTTGTTTCAGATCGGTTGCATTGGCTTGCTC 210
                |  |  ||    ||  ||      |       |||  |||  ||||     ||     ||  |||||||  |
Db          448 TATGTCGGTCGTGGTATGCAGTTCCTTGACTTGATTCAAGAAGGAAATATGGGCTTGATG 507

Qy          211 AAGGCCGTTGACAAGTTCGATCTTTCGTACGATGTGAGATTTTCGACCTATGCGGTGCCA 270
                |||||  ||||||||||||  ||    |||  |  |    |  |   |||||  ||  |||||
Db          508 AAGGCGGTTGACAAGTTTGACTATTCTAAAGGGTTCAAGTTTTCAACTTATGCAACTTGG 567

Qy          271 ATGATCATCGGAGAAATTCAACGCTTTTTGCGCGATGACG---GTACGGTTAAGGTCAGT 327
                 |||        |   ||     ||     |  |     ||   |  |    ||||   |       ||
Db          568 TGGATTCGTCAGGCTATCACTCGTGCTATTGCAGACCAAGCTCGTACCATCCGTATCCCA 627

Qy          328 CGATCGTTAAAAGAAACAGCGAATAAGGTGCGGCGATCAAAGGATGAATTGTACAAGCAA 387
                    |     |||||     |||||   ||    ||    |  ||     |   |     |  ||
Db          628 GTTCACATGGTTGAAACTATCAATAAATTGGTTCGTGAACAGCGGAATCTCCTTCAAGAA 687

Qy          388 TTCGGCCGTGCCCCCACGATCGCAGAAGTGGCAGAAGCAGTGGGAATCACGCCGGAGGAA 447
                ||  ||  |   |  ||  ||     |  |  |  |  |||   |  |||  ||  ||  ||  |
Db          688 TTGGGGCAAGATCCGACACCAGAACAGATTGCTGAACGAATGGATATGACACCTGATAAG 747

Qy          448 GT 449
                ||
Db          748 GT 749
```

```
RESULT 14
AEL12413
ID   AEL12413 standard; DNA; 915 BP.
XX
AC   AEL12413;
XX
DT   28-DEC-2006 (first entry)
XX
DE   Streptococcus pneumoniae strain 14453 protein-encoding DNA, SEQ:1096.
XX
KW   Vaccine; diagnosis; drug discovery; protein production;
KW   bacterial infection; Streptococcus pneumoniae infection;
KW   bacterial meningitis; bacterial pneumonia; bacteremia; otitis media;
KW   antibacterial; neuroprotective; antiinflammatory; respiratory-gen.;
KW   auditory; gene; ds.
XX
OS   Streptococcus pneumoniae; strain 14453.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..915
FT                   /*tag= a
FT                   /product= "Streptococcus pneumoniae protein SEQ ID
FT                   NO:3757"
XX
PN   US7122368-B1.
XX
PD   17-OCT-2006.
XX
PF   30-DEC-2004; 2004US-00027877.
XX
PR   02-JUL-1997;    97US-0051553P.
PR   12-MAY-1998;    98US-0085131P.
PR   30-JUN-1998;    98US-00107433.
PR   26-MAY-2000; 2000US-00583110.
PR   14-AUG-2003; 2003US-00640833.
XX
PA   (SNFI ) SANOFI PASTEUR LTD.
XX
PI   Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;
XX
DR   WPI; 2006-812612/82.
DR   P-PSDB; AEL15074.
XX
```

PT New isolated nucleic acid and polypeptide isolated from strain 14453 of
PT Streptococcus pneumoniae, useful for diagnosing, preventing, or treating
PT bacterial infections, e.g. S. pneumoniae infection.
XX
PS Example; SEQ ID NO 1096; 29pp; English.
XX
CC The invention relates to an isolated nucleic acid, especially (AEL12919),
CC which encodes the Streptococcus pneumoniae protein of AEL15580. This
CC nucleic acid is one of 2661 disclosed protein-encoding nucleic acids
CC (AEL11318-AEL13978) isolated from a Streptococcus pneumoniae strain 14453
CC genomic library whose predicted products (AEL13979-AEL16639) exhibit
CC homology to known prokaryotic, eukaryotic or archaeal open reading frames
CC (ORFs) or proteins. The invention also relates to a recombinant
CC expression vector comprising the nucleic acid of the invention operably
CC linked to a transcription regulatory element; and a host cell comprising
CC the recombinant expression vector. The Streptococcus pneumoniae nucleic
CC acids and proteins of the invention are useful for diagnosing,
CC preventing, or treating pathological conditions resulting from bacterial
CC infections, especially infections caused by Streptococcus pneumoniae such
CC as meningitis, bacteremia, pneumonia and otitis media. They may also be
CC used in vaccine compositions for the treatment of Streptococcus
CC pneumoniae infections and as targets for antibacterial drugs.
CC Additionally the nucleic acids are useful in the production of
CC commercially important proteins such as enzymes for use in fermentation
CC reactions or in the production of commercially useful metabolites. The
CC present sequence represents a Streptococcus pneumoniae strain 14453
CC protein-encoding nucleic acid which was identified in the exemplification
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=7122368B1.
XX
SQ Sequence 915 BP; 279 A; 175 C; 220 G; 241 T; 0 U; 0 Other;

Query Match 10.1%; Score 76.4; DB 21; Length 915;
Best Local Similarity 50.5%; Pred. No. 2.4e-13;
Matches 213; Conservative 0; Mismatches 206; Indels 3; Gaps 1;

```
Qy          31 CCATTTCTGACCAATGACCAAGTGAAAGATTTGATAGCCAAGAGCCAAGCTGGCGATACG 90
               ||  |  ||||||||||  | | ||||||| |||  |           |||||||| ||||
Db         328 CCTCTCTTGACCAATGAAGAGGAGAAAGAGTTGGCACTGGCTGTTGAAGCTGGTGATATC 387

Qy          91 GATGCACGTGAGCTTCTCGTGAATAGCAATATCAGACTGGTCTGGTCCGTCGTCCAGCGC 150
               ||  ||      | | ||| | | |    |||| |  |  ||||     ||| | | | | |||
Db         388 GAAGCCAAACAACGTCTTGCGGAAGCCAATCTTCGTTTGGTTGTTTCCATTGCCAAACGC 447

Qy         151 TTTATCAACCGCGGGTATGAAGCGGATGATTTGTTTCAGATCGGTTGCATTGGCTTGCTC 210
               | | ||   || ||      |      ||| ||| ||||    ||     || ||||||| |
Db         448 TATGTCGGTCGTGGTATGCAGTTCCTTGACTTGATTCAAGAAGGAAATATGGGCTTGATG 507

Qy         211 AAGGCCGTTGACAAGTTCGATCTTTCGTACGATGTGAGATTTTCGACCTATGCGGTGCCA 270
               ||||| |||||||||||| ||   |||  | |   | |  |||||  || |||||
Db         508 AAGGCGGTTGACAAGTTTGACTATTCTAAAGGGTTCAAGTTTTCAACTTATGCAACTTGG 567

Qy         271 ATGATCATCGGAGAAATTCAACGCTTTTTGCGCGATGACG---GTACGGTTAAGGTCAGT 327
                |||      |  ||    ||   | |    ||  | |    ||||  |     ||
Db         568 TGGATTCGTCAGGCTATCACTCGTGCTATTGCAGACCAAGCTCGTACCATCCGTATCCCA 627

Qy         328 CGATCGTTAAAAGAAACAGCGAATAAGGTGCGGCGATCAAAGGATGAATTGTACAAGCAA 387
                   |    |||||    |||||  ||   ||   | ||    |  |     |  ||
Db         628 GTTCACATGGTTGAAACTATCAATAAATTGGTTCGTGAACAGCGGAATCTCCTTCAAGAA 687

Qy         388 TTCGGCCGTGCCCCCACGATCGCAGAAGTGGCAGAAGCAGTGGGAATCACGCCGGAGGAA 447
               ||  || |  |  || ||     | | |  | || |||   | |||  || || || ||  |
Db         688 TTGGGGCAAGATCCGACACCAGAACAGATTGCTGAACGAATGGATATGACACCTGATAAG 747

Qy         448 GT 449
               ||
Db         748 GT 749
```

```
RESULT 15
AEL50821
ID   AEL50821 standard; DNA; 915 BP.
XX
AC   AEL50821;
XX
DT   28-DEC-2006  (first entry)
XX
DE   Streptococcus pneumoniae strain 14453 protein-encoding DNA, SEQ:1096.
XX
KW   Vaccine; diagnosis; drug discovery; protein production;
KW   bacterial infection; Streptococcus pneumoniae infection;
KW   bacterial meningitis; bacterial pneumonia; bacteremia; otitis media;
KW   antibacterial; neuroprotective; antiinflammatory; respiratory-gen.;
KW   auditory; gene; ds.
XX
OS   Streptococcus pneumoniae; strain 14453.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..915
FT                   /*tag= a
FT                   /product= "Streptococcus pneumoniae protein SEQ ID
FT                   NO:3757"
XX
PN   US7129340-B1.
XX
PD   31-OCT-2006.
XX
PF   30-DEC-2004; 2004US-00028457.
XX
PR   02-JUL-1997;    97US-0051553P.
PR   12-MAY-1998;    98US-0085131P.
PR   30-JUN-1998;    98US-00107433.
PR   26-MAY-2000; 2000US-00583110.
PR   14-AUG-2003; 2003US-00640833.
XX
PA   (SNFI ) SANOFI PASTEUR LTD.
XX
PI   Doucette-Stamm L, Bush D, Zeng Q, Opperman T, Houseweart CE;
XX
DR   WPI; 2006-812716/82.
DR   P-PSDB; AEL53482.
XX
PT   New isolated nucleic acid and polypeptide isolated from Streptococcus
PT   pneumoniae, useful as components of antibacterial vaccines, and for
PT   diagnosing or treating S. pneumoniae and other Streptococcus infections.
XX
PS   Example; SEQ ID NO 1096; 29pp; English.
XX
CC   The invention relates to an isolated nucleic acid, especially (AEL52290),
CC   which encodes the Streptococcus pneumoniae protein of AEL54951. This
CC   nucleic acid is one of 2661 disclosed protein-encoding nucleic acids
CC   (AEL49726-AEL52386) isolated from a Streptococcus pneumoniae strain 14453
CC   genomic library whose predicted products (AEL52387-AEL55047) exhibit
CC   homology to known prokaryotic, eukaryotic or archaeal open reading frames
CC   (ORFs) or proteins. The invention also relates to a recombinant
CC   expression vector comprising the nucleic acid of the invention operably
CC   linked to a transcription regulatory element; and a host cell comprising
CC   the recombinant expression vector. The Streptococcus pneumoniae nucleic
CC   acids and proteins of the invention are useful for diagnosing,
CC   preventing, or treating pathological conditions resulting from bacterial
CC   infections, especially infections caused by Streptococcus pneumoniae such
CC   as meningitis, bacteremia, pneumonia and otitis media. They may also be
CC   used in vaccine compositions for the treatment of Streptococcus
CC   pneumoniae infections and as targets for antibacterial drugs.
CC   Additionally the nucleic acids are useful in the production of
CC   commercially important proteins such as enzymes for use in fermentation
```

CC reactions or in the production of commercially useful metabolites. The
CC present sequence represents a Streptococcus pneumoniae strain 14453
CC protein-encoding nucleic acid which was identified in the exemplification
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?DocID=7129340B1.
XX
SQ Sequence 915 BP; 279 A; 175 C; 220 G; 241 T; 0 U; 0 Other;

Query Match 10.1%; Score 76.4; DB 21; Length 915;
Best Local Similarity 50.5%; Pred. No. 2.4e-13;
Matches 213; Conservative 0; Mismatches 206; Indels 3; Gaps 1;

```
Qy          31 CCATTTCTGACCAATGACCAAGTGAAAGATTTGATAGCCAAGAGCCAAGCTGGCGATACG 90
               ||  |  ||||||||||  | | ||||||  |||  |           |||||||  ||||
Db         328 CCTCTCTTGACCAATGAAGAGGAGAAAGAGTTGGCACTGGCTGTTGAAGCTGGTGATATC 387

Qy          91 GATGCACGTGAGCTTCTCGTGAATAGCAATATCAGACTGGTCTGGTCCGTCGTCCAGCGC 150
               || ||      | | ||| | | |    |||| |  |  ||||     ||| | | | | |||
Db         388 GAAGCCAAACAACGTCTTGCGGAAGCCAATCTTCGTTTGGTTGTTTCCATTGCCAAACGC 447

Qy         151 TTTATCAACCGCGGGTATGAAGCGGATGATTTGTTTCAGATCGGTTGCATTGGCTTGCTC 210
               | | ||   || ||      |      ||| ||| ||||    ||     || |||||| |
Db         448 TATGTCGGTCGTGGTATGCAGTTCCTTGACTTGATTCAAGAAGGAAATATGGGCTTGATG 507

Qy         211 AAGGCCGTTGACAAGTTCGATCTTTCGTACGATGTGAGATTTTCGACCTATGCGGTGCCA 270
               ||||| ||||||||||| ||   |||  | |   | |  ||||| || |||||
Db         508 AAGGCGGTTGACAAGTTTGACTATTCTAAAGGGTTCAAGTTTTCAACTTATGCAACTTGG 567

Qy         271 ATGATCATCGGAGAAATTCAACGCTTTTTGCGCGATGACG---GTACGGTTAAGGTCAGT 327
                 |||      |  ||    ||   | |    ||  | |   ||||  |     ||
Db         568 TGGATTCGTCAGGCTATCACTCGTGCTATTGCAGACCAAGCTCGTACCATCCGTATCCCA 627

Qy         328 CGATCGTTAAAAGAAACAGCGAATAAGGTGCGGCGATCAAAGGATGAATTGTACAAGCAA 387
                   |    |||||   |||||  ||   ||   | ||     |  |     |  ||
Db         628 GTTCACATGGTTGAAACTATCAATAAATTGGTTCGTGAACAGCGGAATCTCCTTCAAGAA 687

Qy         388 TTCGGCCGTGCCCCCACGATCGCAGAAGTGGCAGAAGCAGTGGGAATCACGCCGGAGGAA 447
               || || |  |  || ||    | | |  | || ||| | |||  || || || ||  |
Db         688 TTGGGGCAAGATCCGACACCAGAACAGATTGCTGAACGAATGGATATGACACCTGATAAG 747

Qy         448 GT 449
               ||
Db         748 GT 749
```

Search completed: November 4, 2008, 17:14:39
Job time : 247 secs

SCORE 3.0